MIT.8924

UNITED STATES PATENT APPLICATION

Of

RAHUL SARPESHKAR

And

LORENZO TURICCHIA

For a

SYSTEM AND METHOD FOR DISTRIBUTED GAIN CONTROL

PRIORITY

The present application claims priority to U.S. Provisional Patent Application Ser. No. 60/398,253 filed July 24, 2002.

BACKGROUND

The invention generally relates to spectral enhancement systems for enhancing a spectrum of multi-frequency signals (e.g., acoustic, electromagnetic, etc.), and relates in particular to spectral enhancement systems that involve filtering and amplification.

Conventional spectral enhancement systems typically involve filtering a complex multi-frequency signal to remove signals of undesired frequency bands, and then amplifying the filtered signal in an effort to obtain a spectrally enhanced signal that is relatively background free.

In many systems, however, the background information may be difficult to filter out based on frequencies alone because the complex multi-frequency signal may include background noise that is close to the frequencies of the desired information signal. Moreover, many conventional spectral enhancement systems inadvertently amplify some background noise with the amplification of the desired information signal.

For example, a spectral enhancement system may include one or more band pass filters into which an input signal is received, as well as one or more compression and/or amplification units, the outputs of which are combined at a combiner to produce an output signal. If the frequencies of the desired signals, for example, vowel sounds in an auditory signal are either within a band filtered frequency or are surrounded by

substantial noise signals in the frequency spectrum, then such a filter and amplification system may not be sufficient in certain applications.

As a particular example of a spectral enhancement system, an electronic cochlea models the traveling-wave amplifier architecture of a biological cochlea as a cascade of nonlinear-and-adaptive second-order filters with corner frequencies that decrease exponentially from approximately 10kHz to 100Hz. Due to the successive compounding of gains, a change in the individual filter gains of a few percent can alter the gain of the composite transfer function by many orders of magnitude. For example, $(1.1)^{45} = 73$ while $(0.9)^{45} = 0.009$. It is very difficult to accomplish such wide-dynamic-range gain control with one localized amplifier without changing the amplifier's bandwidth, temporal resolution, and power dissipation drastically. Any parameter variations in the Q's of the various cochlear filters, which can result in inhomogenities and nonrobust or unstable operation, are compensated for through gain control. Any physical biological system, such as the cochlea, must possess a feedback system to ensure that it works in a real-world environment, where parameters are not perfectly matched and controlled to high precision as in current digital implementations or simulations.

Distributed gain control and the traveling-wave phenomena are important aspects of the silicon cochlea (as disclosed in *A Low-Power Wide-Dynamic Range Analog VLSI Cochlea*, Sarpeshkar, R., Lyon, R.F., and Mead, C.A., Analog Integrated Circuits and Signal Processing (1998), the disclosure of which is hereby incorporated by reference) in replicating the performance of the biological cochlea. The silicon cochlea's importance to cochlear implant processing is significant for at least the following reasons.

1) An exponentially tapering filter-cascade architecture provides an extremely efficient mechanism for constructing a bank of closely spaced high-order filters as disclosed in *Traveling Waves versus Bandpass Filters: The Silicon and Biological Cochlea*, Sarpeshkar, R., Proceedings of the International Symposium on Recent Developments in Auditory Mechanics, World Scientific (2000), and *Filter Cascades as Analogs of the Cochlea*, Lyon, R.F., Neuromorphic Systems Engineering (1998), the disclosures of which are both hereby incorporated by reference. As the number of channels in implants continues to grow (e.g., 31 channel implants, 64-channel implants, 128-channel implants etc.), the advantages of filter cascades in creating a bank of high-order filters will become more and more apparent.

2) A sophisticated frequency-dependent version of the gain control algorithms presently used in implants and hearing aids may be implemented as disclosed in *Comparison of Different Forms of Compression in Wearable Digital Hearing Aids*, Stone, M.A., Moore, B.C.J., Alcantara, J.I., and Glasberg, B.R., J. Acoustic Society of America, (1999), the disclosure of which is hereby incorporated by reference. Thus loud sounds at one frequency do not have to result in inaudible sounds at another frequency. Also, the gain control allows important phenomena in the perception of speech in noise such as forward masking to be easily modeled. Gain control has been shown to be particularly important in the performance of speech recognition systems in reverberant and noisy environments.

3) The architecture of the cochlea is amenable to both time and place coding as described in *A Low-Power Analog Front-end Module for Cochlear Implants*, Wang,

R.J.W, Sarpeshkar, R, Jabri, M. and Mead, C. XVI World Congress on Otorhinolaryngology (1997), the disclosure of which is hereby incorporated by reference.

4) The biological realism allows several important phenomena in biology to be naturally replicated. These include filter broadening with level, the distributed coding of loudness, the transition from place cues to time cues as level increases, redundant signal representations, the close intertwining of both filtering and compression rather than the artificial separation of filtering and compression in today's implants, compression of long-term information while preserving good sensitivity to transients, two-tone suppression, the upward spread of masking, and forward masking. Although it is quite possible that none of these effects have any importance for implant patients, given that cochlear front ends have been shown to improve speech recognition in noise it is unlikely that models closer to the biology will have no impact on implant patients. It is also likely that coding strategies that are closer to the biology will prove superior to those that are not.

5) The silicon cochlea's analog circuit techniques provide a foundation for ultra-low-power cochlear implant design.

The silicon cochlea may be implemented as a particular form of local feedforward gain control as disclosed in *A Low-Power Wide-Dynamic Range Analog VLSI Cochlea* discussed above. Such an implementation, however, generates input-output curves that are too compressive as compared with those in a real cochlea. Such curves are not suitable for direct use in cochlear implants. Furthermore, such curves cannot easily be programmed to implement a desired compression characteristic, an important necessity in a practical system.

There is a need therefore, for an improved spectral enhancement system that is efficient and practical.

SUMMARY OF THE ILLUSTRATED EMBODIMENTS

In accordance with an embodiment, the invention provides a spectral enhancement system that includes a plurality of distributed filters, a plurality of energy distribution units, and a weighted-averaging unit. At least one of the distributed filters receives a multi-frequency input signal. Each of the plurality of energy-detection units is coupled to an output of at least one filter and provides an energy-detection output signal. The weighted-averaging unit is coupled to each of the energy-detection units and provides a weighted-averaging signal to each of the filters responsive to the energy-detection output signals from each of the energy-detection units to implement distributed gain control. In an embodiment, the energy detection units are coupled to the outputs of the filters via a plurality of differentiator units.

BRIEF DESCRIPTION OF THE DRAWINGS

The following description may be further understood with reference to the accompanying drawings in which:

Figure 1 shows an illustrative diagrammatic schematic view of a portion of a system in accordance with an embodiment of the invention;

Figures 2A - 2C show illustrative diagrammatic graphical views of spatial kernels for implementing distributed gain control in accordance with systems of various embodiments of the invention;

Figure 3 shows an illustrative diagrammatic graphical view of response characteristics of systems of various embodiments of the invention at various amplitudes for single tone stimulations;

Figures 4A and 4B show illustrative diagrammatic graphical views of input-output transfer functions for different values of the power law of the compression characteristic;

Figures 5A and 5B show illustrative diagrammatic graphical views of spatial responses for two-tone stimulations for different frequencies of the non-dominant tones;

Figure 6A shows an illustrative diagrammatic graphical view of a sample spectrum of the phoneme /u/; and

Figures 6B - 6C show illustrative diagrammatic graphical views of spatial response profiles for the sample of Figure 6A with and without gain control.

The drawings are shown for illustrative purposes only and are not to scale.

DETAILED DESCRIPTION OF THE ILLUSTRATED EMBODIMENTS

It has been discovered that a system may be developed to provide an efficient spectral enhancement system by employing a bank of wide-dynamic-range frequency-analysis channels. Such a system may be created using hardware circuit components (e.g., electronic, optic or pneumatic), using software, or using any other simulation routine such as the MATLAB program sold by Math Works, Inc. of Natick, Massachusetts.

For example, in an auditory enhancement or replacement systems for humans, an electronic cochlea maps the traveling-wave architecture of the biological cochlea into a

silicon chip. In both biology and electronics gain control is essential in ensuring that the architecture is robust to parameter changes, and in attaining wide dynamic range. A silicon cochlea with distributed gain control is advantageous as a front end in cochlear-implant processors to improve patient performance in noise and to implement the computationally intensive algorithms of the biological cochlea with very low power.

In accordance with an embodiment, the invention provides a computer simulation of a filter-cascade cochlear model with distributed gain control that incorporates several important features such as multi-band compression, an intertwining of filtering and compression, masking, and an ability to tradeoff the preservation of spectral contrast with the preservation of audibility. The gain control algorithm disclosed herein successfully reproduces cochlear frequency response curves, and represents an example of a class of distributed-control algorithms that could yield similar results. In distributed gain-control systems like the cochlea, each individual filter does not change its gain appreciably although the collective system does change its gain appreciably. Thus, a system may maintain its bandwidth, temporal resolution, and power dissipation to be relatively invariant with amplitude.

Figure 1 shows a schematic architecture 10 for implementing a distributed-gain-control system in a silicon cochlea in accordance with an embodiment of the invention. In certain embodiments, it is desired to obtain a gain-control strategy that functions well for use in cochlear-implant processors. The system is shown for a single second order section h_j (18) with the neighboring second order sections being designated h_{j-1} (16), h_{j-2} (14), h_{j-3} (12), h_{j+1} (20), h_{j+2} (22), h_{j+3} (24). The output signals from each the sections 12 - 24 are optionally coupled to a plurality of differentiators 26 - 36 as shown and provided

to a plurality of independent energy detection units 38 - 48. The outputs of the energy-detection units 38 - 48 are coupled to a weighted averaging kernel 50, and the kernel 50 provides a weighted averaging signal I_j to a non-linearity unit 52, which in turn provides a Q_j signal to the second order section h_j. Each of the second order sections 12 - 24, therefore, is provided a Q signal that is generated by the kernel 50 and non-linearity unit 52 to be responsive to energy-detection signals from each of the sections 12 - 24. The sections 12 - 24 each generally perform a filtering function, and may for example, provide a low pass, band pass or high pass filter function.

During operation, the cascaded resonant second-order sections 12 - 24 may provide low pass filter functions and have characteristic frequencies (CF_j) that are exponentially tapered from the beginning of the cascade to the end of the cascade. The outputs from the resonant low pass second-order sections 12 - 24 are double differentiated in the $(jw/CF_j)^2$ blocks 26 - 36 to create CF-normalized bandpass frequency-response characteristics at each stage of the silicon cochlea. The envelope energy in each of these stages is extracted by the envelope-detector (ED) blocks 38 - 48 and fed to a kernel that computes a spatially-filtered version of these energies. The kernel 50 weights local energies more strongly than energies from remote stages. The output of the kernel, I_j, is then passed through nonlinear block, NL_j (52). The NL block outputs a large value for the resonant gain, Q_j, if the energy is low, and a small value for Q_j, if the energy is high, thus, performing gain control. The attack and release dynamics of the gain control arise from charging and discharging time constants in the envelope detector respectively, and may be tapered with the CF's of the cochlear stages. For clarity, the architecture is only shown in detail for stage j of the cascade, but every stage

of the cascade has similar NL_j blocks that operate on local estimates of envelope energy output by the kernel.

The weighted-averaging at any local filter is a function of the of the energy outputs of each of the other filters as well as the local energy output and may be generally represented as follows:

$$I_j = F_j(\ldots, e_{j-3}, e_{j-2}, e_{j-1}, e_j, e_{j+1}, e_{j+2}, e_{j+3}, \ldots) \qquad (1)$$

A specific example of the equations that describe distributed gain control in accordance with an embodiment of the invention are disclosed in the equations below to describe the spatial weight-averaging kernel and non-linearity unit:

Spatial kernel: $I_j = \sum_{i=1}^{N} w_i^j e_i$ (2)

NL: $Q_j = Q_{\max}$ for $I_j \le K$ and

$$Q_j = \frac{(Q_{\max} - Q_{\min})}{(I_j / K)^z} + Q_{\min} \text{ for } I_j > K \qquad (3)$$

The weights of the kernel are given by w_i^j. The parameters $Q_{\max}$ and $Q_{\min}$ determine the maximum and minimum Q settings of a cochlear stage. The value K determines the knee of the cochlear compression characteristic, and z determines the power law of the compression characteristic. A large K implies that the gain control is activated only at large intensities. A large z means that the compression characteristic obeys a small power law, and is relatively flat with intensity. The spatial extent of the

kernel, Q_{max}, and Q_{min} determine whether the gain control is broadband and preserves spectral contrast (large spatial-extent kernels and small Q's) or whether it is narrowband and preserves audibility (small spatial-extent kernels and large Q's).

Figures 2A - 2C show three examples of kernels for use in various embodiments of the invention. The kernels are shown for the Q control of stage 60. The kernel shown at 54 in Figure 2A, labeled K_{-1} is a purely feedforward kernel with gain control inputs arising from only the stage previous to that being controlled. The kernel shown at 56 in Figure 2B, labeled K_{hoct}, has inputs to the gain control arising from only stages a half octave ahead of the stage being controlled. The kernel shown at 58 in Figure 2C is a purely feedback kernel. Stages that are a one-half octave ahead are the most strongly affected by the local stage's gain always, independent of the gain control. The kernel shown in Figure 2C, labeled K_{exp}, has exponential weighting for stages beyond a one-half octave and before a one-half octave of the stage being controlled. Each of the kernels has various pros and cons. K_{-1} is simple and fast and has no stability issues. The kernel K_{hoct} may result in instability in the gain control if the adaptation time constants are too fast. A cascade architecture that incorporates complex zeros to reduce the group delay in the second order sections may help improve the stability and speed-of adaptation tradeoff in schemes using K_{hoct}. The kernel K_{exp} behaves similar to K_{hoct} but the resulting gain control and masking are more broadband. Interesting results may be obtained for a K_{-1} kernel using MATLAB simulations.

As shown in Figure 3, the cochlear frequency response curves at various intensities (1.1 dB, 20 dB, 40 dB, 60 dB and 80 dB) are shown (at 60, 62, 64, 66, and 68 respectively). Figure 3 shows pure-tone cochlear response characteristics at various

intensities for a cascade with 24 filters per octave, a Q_{min} of 0.7, a Q_{max} of 1.2, a K_{-1} kernel, and parameters of $K = 1000$ and $z = 0.4$.. The adaptation and broadening in resonant gain, compression, and peak shifts are all evident. Figure 3 shows that in response to a pure tone at various intensities, 1) the peak is broadened, 2) the peaks are compressed, and 3) the peaks shift to the left as the signal intensity is increased.

Input-output curves are shown in Figures 4A for output = input, output = A*input$^{0.18}$, output = Coch Resp with z=0.2 in Equation (3) above, output = Coch Resp with .z=0.4, and output = Coch Resp with z=0.8 at 70, 72, 74, 76 and 78 respectively. Figure 4A shows that as z is varied, the power law of the compression characteristic at the best frequency (BF) may be changed. Figure 4(B) shows that as we vary K, the knee of the compression characteristic at the best frequency is changed. The input-output curves for output = input, output = A*input$^{0.18}$, output = Coch Resp (K=1e2), output = Coch Resp (K=1e3), and output = Coch Resp (K=1e4) are shown in Figure 4B at 80, 82, 84, 86 and 88 respectively. Figure 4B shows a compression characteristic of an algorithm in accordance with an embodiment of the invention

Figures 5A and 5B shows the cochlear spatial responses 90 and 92 respectively for a two-tone stimulation as the frequency of the nondominant tone is varied with respect to the dominant tone. Figure 5A shows the masking phenomena for two-tone stimulation due to gain control for a K_{-1} kernel, and Figure 5B shows the masking phenomena for two-tone stimulation due to gain control for a K_{exp} kernel. These figures demonstrate that the model in accordance with an embodiment performs a two-tone suppression with a winner-take-all behavior (i.e., the smaller of the two tones is suppressed).

Figures 6B - 6C show cochlear spatial response profiles with and without gain control for the multi-frequency signal shown in Figure 6A. Figure 6A shows at 94 the multi-frequency signal for the phoneme /u/. Figure 6B shows the spatial response profile 96 of the cochlea when the input is the phoneme /u/ without gain control. Figure 6C shows the spatial response profile 98 of the cochlea when the input is the phoneme /u/ with gain control. The gain control ensures that all three formants are important in discrimination. As shown in Figure 6A, the signal 94 includes three distinct peaks F1, F2 and F3 that vary in intensity. When the gain control is on, the three peaks are all at a similar level (equalization) as shown at 98 in Figure 6C. By comparison, when the gain control is off, the distinctiveness of the peaks F2 and F3 is largely lost in the signal 96 as shown in Figure 6B.

Those skilled in the art will appreciate that numerous variations and modifications may be made to the above embodiments without departing from the spirit and scope of the claims.

What is claimed is: